

WORKING PAPER SERIES



Marco Li Calzi, Paolo Pellizzari

Simple Market Protocols
for Efficient Risk Sharing

Working Paper n. 136/2006
July 2006

ISSN: 1828-6887

This Working Paper is published under the auspices of the Department of Applied Mathematics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional nature.

Simple Market Protocols for Efficient Risk Sharing*

MARCO LICALZI

<licalzi@unive.it>

Dept. of Applied Mathematics

University of Venice and SSAV

PAOLO PELLIZZARI

<paolop@unive.it>

Dept. of Applied Mathematics

University of Venice and SSAV

(July 2006)

Abstract. This paper studies the performance of four market protocols with regard to allocative efficiency and other performance criteria such as volume or volatility. We examine batch auctions, continuous double auctions, specialist dealerships, and a hybrid of these last two. All protocols are practically implementable because the messages that traders need to use are simple. We test the protocols by running (computerized) experiments in an environment that controls for traders' behavior and rules out any informational effect. We find that all protocols generically converge to the efficient allocation in finite time. An extended comparison over other performance criteria produces no clear winner, but the presence of a specialist is associated with the best all-round performance.

Keywords: market microstructure, allocative efficiency, comparison of market institutions, performance criteria.

JEL Classification Numbers: G19, D61, D44, C63.

Correspondence to:

Marco LiCalzi Dept. of Applied Mathematics, University of Venice
Dorsoduro 3825/e
30123 Venezia, Italy

Phone: [++39] (041)-234-6925

Fax: [++39] (041)-522-1756

E-mail: licalzi@unive.it

* We thank Alessandro Sbuelz, an anonymous referee, and audiences at the universities of Bocconi, Cattolica Milano, Palermo, and Salerno as well as SIAM 2005 Annual Meeting for their comments. Financial support from MIUR is acknowledged.

1 Introduction

Financial markets where agents exchange risky assets serve two main purposes. First, they allocate risk among traders and improve allocative efficiency. Second, they diffuse traders' private information and facilitate information diffusion. The simultaneous pursuit of allocative and informational efficiency is usually impossible. Different market arrangements are more favorable to the search for different notions of efficiency. We observe that the state of knowledge in this respect is remarkably unbalanced.

There is a vast literature on market microstructure that is especially keen on the analysis of the conditions affecting the revelation (and exploitation) of private information. On the other hand, much less attention has been devoted to the functioning of financial markets with respect to allocative efficiency. This problem is the focus of our paper, which aims to provide the experimental evidence needed to ground a theoretical analysis.

We study the performance of four market protocols with regard to allocative efficiency and other performance criteria such as volume or volatility. These additional criteria are usually extolled by exchange regulators because they can be objectively measured and provide useful proxies for the evaluation of a market protocol. The four market protocols that we examine are: the batch auction, the continuous double auction, a special form of (nondiscretionary) specialist dealership, and a hybrid of these last two. Contrary to theoretical constructs such as Walrasian tâtonnement, these four protocols are practically implementable because the messages that traders need to use are simple.

We test the protocols by running (computerized) experiments in an environment that controls for traders' behavior and rules out any informational effect. The behavior of the agents span how they formulate trading strategies, how they form expectations, and how they interpret signals. Working with agent-based simulations instead of human agents permits to isolate the impact of the trading protocols from these behavioral components. In standard laboratory experiments, instead, it is not possible to extricate the interactions between protocol and behavioral effects.

The main behavioral limitation on our agents is that they exhibit limited intelligence, similarly to the “zero intelligence” traders in Gode and Sunder (1993). The price of the risky asset is the main driver for their choices; but, like real traders in real markets, they ignore the correct equilibrium price and thus lack an essential piece of information to compute the efficient allocation. This leaves the market protocol in charge of “discovering” the right price for them. From a roaring and confused crowd of traders each trying to (guess and) achieve his preferred risk allocation, the market protocol must extract and send out price signals that point traders in the right direction. This makes convergence to the “right” price

a necessary condition for allocative efficiency. Assuming that there is sufficient liquidity in the market, we find that all protocols generically converge to the efficient allocation and to the equilibrium price in finite time. It is worth noting that our protocols share several characteristics with the general stochastic decentralized resource allocation process developed in Hurwicz, Radner and Reiter (1975), which gives a formal proof for its convergence in finite time.

We then turn to a dynamic analysis of the performances. For practical purposes, it is probably more important to know how protocols perform during the (perhaps long) transient period before they achieve the efficient allocation. We study how long it takes for different protocols to discover the equilibrium price and how fast they lead traders to the efficient risk sharing allocation. We measure the volume of trade developed to reach the efficient allocation from the initial endowment, as well as the volatility of the time series of prices. This extended comparison over several dynamic performance criteria produces no clear winner, but the presence of a nondiscretionary specialist dealer is associated with the best all-round performance. Continuing the analogy above, the introduction of an nondiscretionary “center” in a market protocol seems to improve its ability to stabilize and direct traders’ own groping for the right price.

The organization of the paper is the following. Section 2 describes the model tested in our experiments. In particular, Section 2.5 provides a comparative review of our model against the relevant literature. Section 3 details the experimental design and provides detailed instructions for its replication. Section 4 reports on the results obtained and Section 5 offers our conclusions.

2 The model

Following Smith (1982), we identify three distinct components for our (simulated) exchange markets. The environment in Section 2.1 describes the general characteristics of the economy, including agents’ preferences and endowments. The market protocols in Section 2.2 provide the institutional details which regulate the functioning of an exchange. The behavioral assumptions in Section 2.3 specify how agents make decisions and take actions. Section 2.4 details a few alternative behavioral assumptions that have been used to test the robustness of our conclusions. Finally, Section 2.5 compares how our assumptions fares with those prevailing in the literature.

2.1 The environment

We consider a two-asset economy with n traders. The two available assets are a risky stock and cash. The rate of interest is normalized to zero, so cash acts as the numeraire. The stock pays no dividends and has a (random) realization value Y at a given time T in the far future. Each trader i has an initial endowment of cash $c_i \geq 0$ and shares $s_i \geq 0$. The total amount of cash and stock in the economy is $C = \sum_i c_i > 0$ and $S = \sum_i s_i > 0$.

To rule out any informational effect, we assume that all traders believe that Y is normally distributed with mean $\mu \geq 0$ and precision $\tau = 1/\sigma^2 > 0$ and that no new information is ever released. Therefore, traders' beliefs about Y are homogeneous and never change until uncertainty resolves.

Each trader i has CARA preferences over his final wealth, with a coefficient of risk tolerance $k_i > 0$. Therefore, trader i 's excess demand function for stock (net of his endowment s_i) is the linear function

$$q_i(p) = k_i \tau (\mu - p) - s_i. \quad (1)$$

Let $K = \sum_i k_i$ be the sum of traders' coefficients of risk tolerance. By well-known results pioneered in Wilson (1968), the unique efficient risk-sharing allocation for this economy requires that trader i holds $s_i^* = (S/K)k_i$ shares of the stock. In other words, the efficient allocation is unique and proportional to the coefficients of risk tolerance.

2.2 The market protocols

Clearly, the competitive equilibrium achieves the efficient allocation in this environment. The zero aggregate excess demand condition implies

$$p^* = \mu - \frac{S}{\tau K}. \quad (2)$$

At price p^* , the trader i 's net demand

$$q_i(p^*) = \left(\frac{S}{K} \right) k_i - s_i$$

is exactly filled, making his final allocation $q_i(p^*) + s_i$ equal to the required $s_i^* = (S/K)k_i$. Hence, if a market protocol attains the competitive equilibrium, it implements the efficient allocation.

The issue, however, is that the informational requirements for a competitive equilibrium are often not realistic. For instance, the fictitious protocol of the Walrasian auctioneer requires an iterative process during which traders communicate their entire excess demand function to a centralized market maker before any trade takes actually place. Realistic

market protocols are much simpler, in the sense that they require far less information from traders.

We compare the performances of four *simple* market protocols: a batch auction, a continuous double auction, an automated dealership, and a hybrid market. The first protocol is simultaneous, while the other three are sequential. Except where otherwise noted, the following features are common to all the four protocols.

A protocol is organized in trading sessions (or days). Agents participate to every trading session, but each of them can exchange at most one unit in each trading day. During a trading session, an agent can buy or sell at most one unit of the risky asset. If the protocol is sequential, the order in which agents place their orders is randomly chosen for each trading session. If the protocol is simultaneous, all order are made known and processed simultaneously so the time of their submission is irrelevant. In every trading session, each agent selects randomly one side of the market where he attempts to place a trade: he can switch roles across trading sessions, but he cannot place simultaneous orders for buying and selling within the same session. The books are completely cleared at the end of each trading session.

Prices are quoted using a minimum tick; in other words, they are discretized throughout the paper. Moreover, prices must be nonnegative: if a trader places a bid lower than zero, this is ignored; if a trader places an ask lower than zero, this is automatically converted to the lowest strictly positive price compatible with the existing tick.

Batch auction. In each trading session, after traders submit their orders, the aggregate excess demand function is computed and the exchange price p is determined by setting the aggregate excess demand equal to zero. If there are multiple solutions, we select the midpoint of the interval between the lowest and the highest clearing price. (If there are no solutions, no exchange takes place.) Shares and corresponding payments are exchanged between traders who submitted bids no lower than p and asks no higher than p . Traders who placed orders exactly at price p may be accordingly rationed. This protocol is also known as the k -double auction, with $k = 1/2$.

Continuous double auction. In each trading session, traders place their orders on the selling and buying books. Their orders are immediately executed if they are marketable; otherwise, they are recorded on the books with the usual price-time priority. Orders are canceled only when a matching order arrives or the trading day is over.

Automated dealership. There is a specialist dealer who posts bids and asks valid only for a unit transaction. Agents check sequentially the dealer's quotes for the side of the transaction they are attempting. If an agent accepts the dealer's quote, the exchange takes place at the quoted price. Right after a transaction is completed, the two dealer's quotes for bid and ask increase by one tick if the agent completed a purchase and decrease by one tick otherwise. The size of the bid-ask spread stays fixed over time, so the price is never unique. Limited to this protocol, therefore, convergence of prices to a given value p should be interpreted as convergence of prices to a bid-ask interval that contains p . Throughout the paper, we maintain a feminine gender for the dealer and keep her distinct from the traders who are assumed to be male.

Hybrid market. This protocol combines the continuous double auction with the automated dealership. Distinct selling and buying books hold quotes from the specialist dealer and from the public, respectively. The dealer must post bids and asks valid only for a unit transaction and revises her quotes as in the automated dealership; in particular, she moves her quotes only after transactions in which she has been involved. Agents check sequentially the books for the side of the transaction they are attempting. Their orders are immediately executed at the best price available (which may be different from the specialist's) if they are marketable; otherwise, they are recorded on the traders' book with the usual price-time priority. Agents' orders are canceled only when a matching order arrives or the trading day is over. Hence, once deposited on the traders' book, an order from an agent cannot be executed with the dealer.

We note two limitations to the realism of our assumptions about the market protocols. First, in a sequential protocol, the order of arrival for agents is randomly chosen: this mutes any issue concerning the trade-off between efficacy and immediacy. Second, we assume that agents can trade at most one unit per trading session: this circumvents the problem of choosing the order size. Similar restrictive assumptions are common in the literature; see for instance Glosten and Milgrom (1985), which has inspired our version of specialist dealership.

2.3 Behavioral assumptions

A major obstacle in the study of microeconomic systems is that their performance is jointly determined by the interactions of traders' behavior within the market protocol. As traders may react differently to different market protocols, it is difficult to separate the intrinsic

characteristics of a market protocol from the properties induced by the traders' strategies. Our approach concentrates on the institutional characteristics of the protocols, by making general-purpose assumptions on traders' behavior that constrain their freedom of choice. Except where otherwise noted, these assumptions hold for all the (computerized) experiments reported in this paper.

There are three established assumptions in the literature. One is that traders are restricted to trade one unit at a time. This restriction on traded quantities simplifies the strategy space and allows direct comparisons with existing theoretical results. The second assumption states that buying orders are constrained by the available cash and selling orders by the available endowment of stock; that is, budget constraints hold. This is consistent with a value-based strategy ("buy low, sell high"), which is a seemingly natural requirement of rationality for traders' behaviors. The third assumption is that each trader has a constant valuation for each unit traded. We maintain the first two assumptions, but relax the third one.

In our setting, the demand function (1) of each trader is strictly decreasing. The assumption of constant unit valuations is naturally generalized by deducing the valuation for the next units to trade from the demand function. If the current endowment of a trader is s_i , we invert his excess demand function $q_i = k_i\tau(\mu - p) - s_i$ and derive his valuation for the next n units to trade as

$$v_i(\pm n; s_i, k_i) = \mu - \frac{s_i \pm n}{k_i\tau}, \quad (3)$$

where the \pm sign denotes whether the attempted trade is a purchase or a sale. Clearly, this implies that the reservation price of each trader depends on the side of the transaction he is entering, on his current endowment s_i , and on the coefficient of risk aversion k_i ; for simplicity, we suppress these last two arguments and just write $v_i(\pm n)$. Note that, when a trader is restricted to a unit trade, his (implicit) bid-ask spread is simply $v_i(-1) - v_i(+1) = 2/(k_i\tau)$.

Given his valuation, a trader must decide which side of the transaction he wants to attempt and (if necessary) what price to offer. We assume that, at the start of a trading session, each trader chooses either side with equal probability. This randomized choice is stochastically independent of previous history, endowment, or any other variable of the model. After the choice of the trading side is made, therefore, a trader "knows" that he is going to be a buyer (or a seller) and that his valuation for the next unit he will attempt to buy (or sell) is $v_i(+1)$ (or $v_i(-1)$) from Equation (3). Once his trading intentions are known, a trader (deterministically) places a bid or an ask equal to his valuation. An agent posting a

price equal to his valuation may still trade at a lower price, but he is “truthfully” revealing his willingness to buy or sell one more unit. Therefore, we nickname this assumption by TT as a mnemonic for “truth-telling” behavior.

The common properties of our market protocols may impose two deviations from truth-telling. First, the valuation of a trader may be a number different from the ticked prices. As detailed below, we adopt an “exact” experimental design that rules out this case. Second, the valuation of a trader may be negative; in this occurrence, all protocols refuse negative bids and automatically update a negative ask to the lowest non-zero ticked price.

2.4 Alternative behavioral assumptions

The set of assumptions in Section 2.3 uniquely determines traders’ behavior in each of the four market protocols examined in this paper. In our simplified environment, therefore, the differences in performances are due only to the institutional differences embodied in the market protocols. This insulation from spurious effects (due to traders’ behavior) makes it possible to evaluate market protocols on their own. Clearly, this insulation carries a cost in realism because it is a sensible assumption that real traders adjust their strategies to the type of market protocol they are forced to use. Rather, the agents in our simulations exhibit zero intelligence under several respects. (See LiCalzi and Pellizzari (2006) for a study of the impact of introducing minimal forms of intelligence.) They do not react to differences in the market protocol; they do not attempt to trade strategically; they do not update their behavior rules over time.

Given that the purpose of the study is to understand the performance of market protocols *per se*, we justify the assumption of zero intelligence because the experimental design should try as much as possible to keep traders’ behavior unchanged across different institutions. On the other hand, since there are several ways to achieve this objective, we have tested the robustness of our conclusions under two alternative scenarios. Their generation is inspired by the observation that trading strategies solve a tradeoff between efficacy and immediacy. Higher asks and lower bids make trading less aggressive. This increases efficacy (conditional on trading taking place) but lowers immediacy. Taking TT as the reference point, we examine how market protocols perform under two alternative set of assumptions that are less and more aggressive, respectively. We nickname them TT− and TT+ with obvious reference to how aggressive are the trading strategies they generate compared to TT. Hereafter is their description. Since traders are restricted to unit trades, we leave it understood for the rest of this section that prices refer to unit trades.

The first set of assumptions is directly inspired by Gode and Sunder (1993). Under the

heading of “zero intelligence”, they assume that a buyer i bids a price uniformly drawn from the interval $[0, v_i(+1)]$ and a seller i asks a price uniformly drawn from the interval $[v_i(-1), M]$, where M is an exogenously given upper bound on the feasible prices. Gode and Sunder (1993) assumes constant valuations, so it seems natural that M be fixed. In our environment, where agent i ’s valuations $v_i(\pm 1) = \mu - (s_i \pm 1)/(k_i \tau)$ are decreasing in the current stock endowment s_i , we prefer to endogenize the choice of intervals from which a trader picks his bid and ask prices.

We assume that a potential buyer i with a current stock endowment s_i bids a price that is uniformly drawn from the ticked prices in the interval $(v_i(+2), v_i(+1)]$. The upper bound of the interval is the same of Gode and Sunder (1993) and implies that an agent never bids a price (valid for one unit) above his valuation for one unit. The lower bound of the interval implies that an agent never bids a price (valid for one unit) so low that he would rather buy two units. Symmetrically, a potential seller i asks a price that is uniformly drawn from the ticked prices in the interval $[v_i(-1), v_i(-2))$. It is worth noting that when these assumptions generate negative prices, they are ignored or updated by the market protocols as described in Section 2.2.

Compared to TT, where each trader i always bids $v_i(+1)$ or asks $v_i(-1)$, this set of assumptions implies that with positive probability bids are lower and asks are higher, so trading under this set of assumptions (nicknamed TT-) is less aggressive. This can be viewed as an approximation for that common form of strategic trading that misrepresents valuations in order to extract more surplus from a transaction. As this makes it harder to complete a transaction, it reduces immediacy; on the other hand, conditional on trading taking place, it increases the probability of a better price and thus improves efficacy.

The second alternative set of behavioral assumptions generates more aggressive trading than TT and is thus nicknamed TT+. It is inspired by a formal symmetry with the less aggressive TT-. Recall that, under TT, a trader i with current endowment s_i bids $v_i(+1)$ or asks $v_i(-1)$. Under TT-, he bids a ticked price from $(v_i(+2), v_i(+1)]$ or asks a ticked price from $[v_i(-1), v_i(-2))$; hence, prices are shifted outward with respect to TT. Under TT+, we will see that a trader bids from within the interval $[v_i(+1), v_i(0))$ or asks from within the interval $(v_i(0), v_i(-1)]$ and thus prices are shifted inward with respect to TT. However, modeling this behavior in an economically plausible way requires some care.

Consider a trader i who has a valuation $v_i(+1)$ for one additional unit. If he buys it at price of p , this exchange generates a *trading surplus* of $v_i(+1) - p$ for him. Similarly, the trading surplus for a seller is $p - v_i(-1)$. Let $g_i(t)$ be the trading surplus realized by trader i at time t . (If he completes no trade, let $g_i(t) = 0$.) Under TT (or TT-) we have

$g_i(t) \geq 0$ for all t ; that is, no trade generates a negative surplus. This is consistent with an individual rationality constraint under which each unit trade must be profitable. In our environment, however, a trader is likely to undergo several rounds of trading before the efficient allocation is attained so that his *cumulative* trading surplus is also of interest.

Define the *cumulative trading surplus* realized by trader i at time t as the (undiscounted) sum $G_i(t) = \sum_{\tau=1}^t g_i(\tau)$ of his past trading surpluses up to time t . If the agent attempts to speed up his trading by using more aggressive pricing than under TT, he may use his cumulative surplus $G_i(t-1)$ from the past trades (up to but not including t) to cover a potential loss in the current unit trade. This turns the individual rationality constraint into the weaker requirement that $G_i(t) \geq 0$ for all t , leaving the sign of $g_i(t)$ unrestricted. That is, even if the agent may occasionally trades at unfavorable terms, his overall trading surplus with respect to the initial position remains positive.

Formally, under TT+ we assume the following. In each trading session t , a potential buyer i with a current stock endowment s_i and a cumulative trading surplus $G_i(t-1)$ bids a price that is uniformly drawn from the ticked prices in the interval $[v_i(+1), a)$, where a is the minimum between $v_i(0)$ and $v_i(+1) + G_i(t-1)$. Symmetrically, a potential seller i asks a price that is uniformly drawn from the ticked prices in the interval $(b, v_i(-1)]$, where b is the maximum between $v_i(0)$ and $v_i(-1) - G_i(t-1)$. Consider the interval $[v_i(+1), a)$ enclosing the possible bid prices. The lower bound is the same price bid under TT so that bids are more aggressive. The upper bound requires the highest possible bid to be smaller than both $v_i(0)$ and $v_i(+1) + G_i(t-1)$: the first inequality makes sure that the bid and the ask associated with i never cross; and the second inequality implies that even in the worst case (when trade occurs at a price of $v_i(+1) + G_i(t-1)$) the cumulative trading surplus for i does not go negative.

2.5 Comparison with the literature

The seminal contribution about computerized experiments in a controlled environment is Gode and Sunder (1993). This paper establishes the allocative efficiency of the continuous double auction under zero intelligence or, in our jargon, under an assumption similar to TT-. Our analysis extends this work into two directions. First, we explicitly compare the allocative efficiency of several different market protocols and find that they all share the ability to lead zero intelligence traders towards the efficient allocation. Second, we replace their assumption of constant unit valuations for each trader with decreasing unit valuations. In particular, this would dispose of the critique in Cliff and Bruten (1997) about possible nonconvergence to the equilibrium price.

Bottazzi, Dosi and Rebesco (2005) compare the allocative efficiency and other performance criteria for three protocols (Walrasian tâtonnement, the batch auction, and the continuous double auction) in an environment different from ours. They consider an ecology of trend followers and noise traders. The trend followers update their expectations about the future price of the equity and all traders can place both market and limit orders for different quantities. Their study is unable to separate behavioral from market effects, leading to the prudent conclusion reported in the abstract: “The results highlight the importance of the institutional setting in shaping the dynamics of the market but also suggest that the latter can become the outcome of a complicated interaction between the trading protocol and the ecology of traders’ behaviors. In particular, we show that market architectures bear a central influence upon the time series properties of market dynamics. Conversely, the revealed allocative efficiency of different market settings is strongly influenced by the trading behavior of agents.” At the end of Section 4.3 we compare our results with theirs, showing that they share a few qualitative features.

Audet, Gravelle, and Young (2002) compare execution quality in a batch auction and in a competitive (and discretionary) dealership market. The paper argues that execution quality is a multidimensional concept, whose appraisal involves some combination of different measures of performance. From a behavioral point of view, the paper assumes liquidity traders subject to informational shocks and derives choices using a complex neural network approximation to Nash behavior. The main conclusion is that (their own variant of) the competitive dealership provides superior execution quality when trading is thin and correlated or when there are large liquidity shocks.

Satterthwaite and Williams (2002) view a market protocol as an algorithm to solve the problem of allocative efficiency. A finite number of agents is privately informed about their valuations. After being exogenously assigned the role of buyers or sellers, these agents are restricted to trade at most one unit in a single trading session. Traders view the market as a one-shot game with incomplete information and play equilibrium strategies. Under these assumptions, allocative efficiency is not attained. However, the paper proves that the batch auction is worst-case asymptotically optimal: as the number of agents grows, it forces the worst-case inefficiency to zero at the fastest possible rate. Transient effects or other measure of performance are not considered.

3 Experimental design

3.1 Identification

A simulation run for our model requires the specification of five global parameters, a list of individual variables for each trader, as well as specific assumptions about market protocol and traders' behavior. The global parameters are the number n of traders, the mean μ and the variance σ^2 of the realization value Y of the stock, the number t of trading sessions, and the size Δ of the tick. Individually, a trader i is characterized by his coefficient k_i of risk tolerance and by his endowment of cash c_i and stock s_i . Finally, for protocols involving the (female) dealer, we need to select her initial quotes.

The market protocols are described in Section 2.2. For ease of reference, we nickname these four protocols as B (batch auction), C (continuous double auction), D (dealership), and H (hybrid market). Similarly, our three sets of behavioral assumptions are described in Section 2.3 and 2.4. They have been nicknamed TT (truth-telling), TT- (less aggressive than TT), and TT+ (more aggressive).

We have run (computerized) simulations for all $4 \times 3 = 12$ possible combinations of protocols and behavioral assumptions, over different instantiations of the parameters. The results reported in Section 4 seem to be robust to substantial changes in the parameters. The only exception that needs a separate study (left to future research) is the case where the overall liquidity $C = \sum_i c_i$ of the system is very low. Therefore, to simplify the presentation, we fix the exemplar parametric configuration reported in Table 1 and report the simulations for the four market protocols under different behavioral assumptions. The initial dealer's quotes are a bid of 745 and an ask of 751, with a fixed bid-ask spread of 6.

	Parameters		Initialization
Global	n	=	1,000
	μ	=	1,000
	σ^2	=	120
	t	=	1,500
	Δ	=	1
Trader	k_i	\in	$\{10, 12, 15, 20, 24, 30, 40\}$
	c_i	=	50,000
	s_i	\in	$\{20, 24, 30, 40, 48, 60, 80\}$

Table 1: Exemplar for identification.

The coefficients of risk aversion k_i and the initial endowments s_i for each trader are obtained as follows. Initially, we generate a vector \mathbf{a} of $n = 1,000$ stochastically independent draws from the uniform distribution over the integers $\{10, 12, 15, 20, 24, 30, 40\}$. Then, keeping \mathbf{a} fixed, we initialize the k_i 's and the s_i 's for a round of simulation by $k_i = \mathbf{a}_{\pi(i)}$ and $s_i = 2 \cdot \mathbf{a}_{\pi'(i)}$, where π and π' are stochastically independent random permutations of the vector \mathbf{a} . This approach has two main advantages.

The first feature concerns the prices posted by traders. Because the tick is $\Delta = 1$, the ticked prices for our exemplar configuration are integers. Recall that buying and selling valuations for n units obey (3). If the right-hand side of (3) is always an integer, under any of our behavioral assumptions traders post bid and ask prices that are integers; and therefore the market protocol need not round them to a ticked price. In the exemplar, we choose integer values for μ and σ^2 and initialize each k_i 's with a divisor of σ^2 . (The choice $\sigma^2 = 120$ makes sure that there is adequate variety among the divisors.) Then, for all the simulations we report in this paper, the prices submitted by the agents are never rounded by the market protocol.

The second related feature affects the competitive equilibrium price given by (2). If the right-hand side of (2) is an integer, the competitive equilibrium price is among the ticked prices that the market protocol can actually generate; and thus it is possible to have *exact* convergence to the equilibrium price supporting the efficient allocation. (Properly speaking, exact convergence is not relevant under the dealership protocol, because the fixed bid-ask spread prevents the price from being unique.) In the exemplar, the vector of s_i 's is a random permutation of the vector of $2k_i$'s. This implies that $S = \sum_i s_i = 2 \sum_i k_i = 2K$. Therefore, by (2), the competitive equilibrium price is $p^* = \mu - 2\sigma^2 = 760$ for those protocols (B and C) where the dealer is absent.¹ Then, for all the simulations of the B and C protocols reported in this paper, exact convergence to the competitive equilibrium price is always attained.

Finally, we stress that the parameters in the exemplar are not an attempt to calibrate the model to any specific set of real data. We do not claim that our model has the descriptive power that would warrant a calibration exercise. Our simulations are blind to any informational effects and thus are not fit to replicate the price dynamics observed in real markets. The purpose of this study is limited to gather evidence on the performance of different market protocols with regard to allocative efficiency, and our results do not

¹ The initial number of shares made available to agents is $S = 43280$ for all simulations reported in this paper. However, when the dealer is present, her trading affects how many shares are available to the agents in the end.

extrapolate to markets subject to informational effects.

3.2 The simulations

A round of testing requires to simulate $4 \times 3 = 12$ possible combinations of protocols and behavioral assumptions. In order to keep experimental conditions as comparable as possible, a typical round of simulations runs as follows. For each batch of 12 combinations, we instantiate parameters according to the exemplar. We reiterate that, although we have tested a large range of exemplars, for simplicity all simulations reported in this paper share the initialization reported in Table 1.

We also try to reduce to a minimum the impact of randomness on the simulations. Under TT, the only sources of randomness are in the order in which agents are sampled under a sequential market protocol and in the choice of which side of the transaction they attempt to complete. Under $TT\pm$, a third source of randomness is in the selection of bids and asks from two intervals of possible choices. For each batch of simulations, we use the same (randomly chosen) sampling and the same (randomly chosen) selection of transacting sides. In other words, the sequence in which players place orders and the side they attempt to transact is the same within each batch of 12 simulations.

At the end of a batch of simulations, we record the time series for prices, volume, and endowments, and compute relevant statistics for the performance criteria discussed in the next section. The simulations have been run using a dedicated package of routines written in Pascal.

3.3 The performance criteria

There are several criteria that can be used to evaluate the performance of a market protocol. Our first focus of interest is the ability of market protocols to attain the efficient allocation, so we report on their convergence to the efficient allocation as well as on the speed with which they achieve it. We measure traded volume, because higher volumes signal less effective protocols that let unnecessary trades take place. The second focus of interest is the dynamic behavior of a protocol (“how does it get there?”) and therefore we also keep track of cumulative volumes and other appropriate indicators over time. For instance, we report standard deviations and kurtosis for the time series of the closing prices. These are useful indicators for assessing the stability of prices over time, even though Stigler (1964) has long ago warned that one cannot take for granted that “smoothness of price movements is the sign of an efficient market” (p. 125). A detailed description of the performance criteria used follows hereafter.

The first basic criterion of allocative efficiency is whether a protocol converges to the efficient allocation in finite time. Convergence to the efficient allocation is generically achieved, at a price equal to the competitive equilibrium price p^* . Clearly, in any of the four protocols the current price may hit p^* before the efficient allocation is attained. Therefore, we evaluate the speed of convergence by recording the number of trading sessions completed before no further trading takes place. For short, we call this number NT as a mnemonic for “no trade”. To evaluate the dynamic properties of the protocols, we also keep track over time of the distance of the current allocation from the efficient one.

Volume is measured by the total number of one-unit transactions completed before attaining convergence to the efficient allocation. We also monitor volume over time, measured by the cumulative number of transactions completed within the first t trading sessions. Under the dealership protocol, the transfer of one unit from a trader to another one must go through the dealer and therefore requires two transactions, instead of just one. Whenever two matching transactions go through the dealer, we record them as one so that the statistics for volume report for each protocol the effective number of transactions completed between agents. This makes volume directly comparable across protocols, even if it fails to record the few units of unmatched inventory that may remain with the specialist.

Another evaluation criterion is the subjective welfare of the traders, which we measure by the certainty equivalent of their current position. The advantage of using certainty equivalents over utilities is that we can compute (and compare) the monetary value of an allocation by summing up the certainty equivalents of all traders. We note that protocols such as the batch auction or the continuous double auction are self-contained, in the sense that traders exchange cash and stock only among themselves. The overall value of an efficient allocation is of course always the same. However, when a round-trip trade goes through a dealer, some money is lost because of the bid-ask spread. Therefore, we expect the overall value of the allocation to be lower for protocols which are not self-contained.

Finally, we keep track of the standard deviations and the kurtosis of the time series of closing prices. (The closing price is the last price at which a transaction has taken place by the end of the current trading session.) For each simulation, we compute these two statistics over the closing prices between the first trading session and the last active trading session; that is, from time $t = 1$ to NT. As usual, we report the standardized kurtosis $\kappa = \mu_4/\mu_2^2 - 3$, where μ_i is the i -th central moment of the empirical distribution. It is well-known that $\kappa = 0$ for a normal distribution. From a dynamic point of view, we report the standard deviations over closing prices computed over a moving time window formed by the last 20 trading sessions.

4 Results

For each of the four market protocols, we have run 100 simulations under the different three behavioral assumptions. To simplify the presentation, we fix the choice of parameters as described above in Table 1 and separately study the variations across the four protocols for each of our three behavioral assumptions. Section 4.1 reports our results under the TT assumption. Section 4.2 provides a similar description under both TT− and TT+. Neither of these two alternative behavioral assumptions entails major qualitative changes in the results, suggesting that these are mainly imputable to differences in the market protocols. A separate robustness analysis in Section 4.3 discusses the choice of parameters different from our exemplar.

From a static point of view, the obvious benchmark for the comparative evaluation of our performance criteria is the Walrasian allocation mechanism known as tâtonnement. While its informational requirements are unwieldy and thus its practical interest is very limited, in our context the tâtonnement protocol is guaranteed to yield the efficient allocation in one (giant) step. This protocol also attains the competitive equilibrium price p^* . Finally, and perhaps more interestingly, it minimizes the number of transactions needed to achieve the efficient allocation because it correctly matches traders with positive excess demand with agents with negative excess demand.

4.1 The TT assumption

Table 2 reports summary statistics computed as averages over 100 different batches of simulations under the TT behavioral assumption. We have checked these averages against the medians for the same data and there are no substantial differences, so that spurious outlier effects are ruled out. Even a qualitative inspection of the data from a sample of individual simulations shows no significant departures from the averages. The table reports several pieces of information for each of the four protocols, as listed in the first column. We comment on each of them.

Prot.	Vol	ExcV	NT	Dist	CE	Loss	SD	Kurt
B	11251.32	2.61%	148.72	0.00	87873.32	0.00%	4.61	6.90
C	28706.28	162%	149.51	0.00	87873.32	0.00%	31.57	3.77
D	11090.54	1.57%	142.97	0.11	87816.42	0.065%	4.16	0.74
H	16141.92	47.2%	145.06	0.012	87844.09	0.033%	2.85	0.016

Table 2: TT: summary statistics for 100 passes with $n = 1000$.

The second and third columns give the traded volume (Vol) and the percentage of excess volume (ExcV) with respect to the number of transactions that the Walrasian protocol would require to achieve in one step the efficient allocation. The batch auction and the specialist dealership generate minimal excess volume: which one performs better in this respect depends on how far are the initial dealer's quotes from the equilibrium price. The continuous double auction is seriously wasteful, while the hybrid protocol sits in between its parents (dealership and continuous double auction). The ranking with respect to volume is $\{\text{Vol}_B, \text{Vol}_D\} < \text{Vol}_H < \text{Vol}_C$, where the notation $\{\text{Vol}_B, \text{Vol}_D\}$ means that the ranking is not conclusive. From a dynamic point of view, Figure 1 confirms this ranking. In particular, note how similar statistics for B and D in Table 2 correspond to overlapping curves in Figure 1.

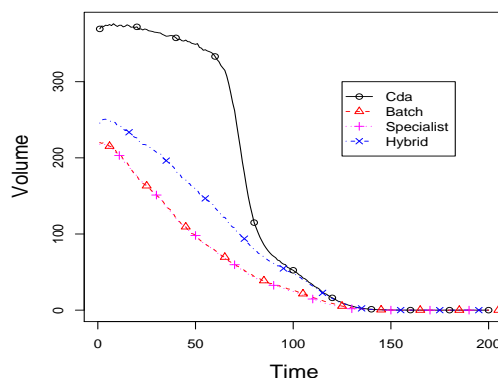


Figure 1: TT: volume versus time.

The fourth column in Table 2 computes the number (NT) of trading days necessary to achieve no trade and hence the efficient allocation. In our exemplar case, the maximum number of units between the initial endowment and the final efficient endowment is 60, so this is a lower bound on the number of trading sessions required to achieve allocative efficiency. The ranking with respect to the time for convergence to no trading is $\text{NT}_D < \text{NT}_H < \text{NT}_B < \text{NT}_C$. The differences are very small, but persistent.

The fifth column reports the distance (Dist) of the final allocation from the efficient allocation. (We measure the distance using the l_1 -norm normalized by the number of traders: thus, $d(e, e^*) = \sum_i |e_i - e_i^*|/n$.) All protocols achieve the efficient allocation, sometimes up to one unit for one trader who cannot find the counterpart for his last unit trade. This problem surfaces only in D and H, because the dealer imposes a spread. This

reduces liquidity in general, and in particular may prevent a trade from occurring when the spread deters an agent from completing a transaction.

From a dynamic point of view, the ability of all protocols to achieve allocative efficiency is shown in Figure 2, which reports the distance of the current allocation from the efficient one with respect to the cumulative volume. (Again, the curves for B and D overlap.) This graph provides an implicit indicator of the ability of a protocol to minimize the number of wasteful trades *during* its search for allocative efficiency. In this respect, it confirms the former ranking $\{\text{Vol}_B, \text{Vol}_D\} < \text{Vol}_H < \text{Vol}_C$ from the perspective of how many trades are needed to achieve a given level of allocative efficiency. In particular, when the efficient allocation is attained, the distance drops down to zero and we recover the values reported in the second column of Table 2.

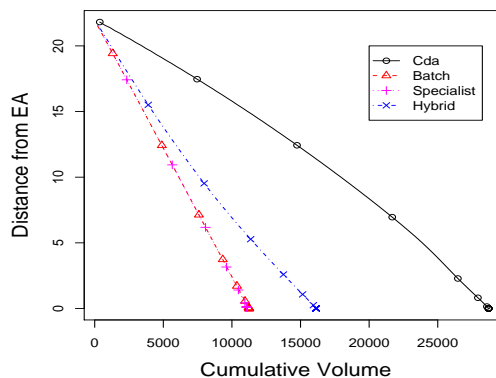


Figure 2: TT: distance from efficient allocation (EA) versus cumulative volume.

Instead of volume, one may also plot the distance from the efficient allocation with respect to time. Since during a trading session all agents are randomly sampled and attempt to make a transaction, this would gauge the ability of a protocol to minimize the number of unnecessary attempted trades on the part of the traders. We do not report the plot here to save space. The ranking over $\{B, D, H\}$ is inconclusive, but each of these protocols consistently requires less time than C to attain a given level of allocative efficiency. Even though the continuous double auction generates higher trade within a session, this does not work towards reducing the number of unnecessary calls in later sessions.

The sixth column in Table 2 reports the arithmetic average of traders' certainty equivalents at time NT, while the seventh column gives the percentage loss with respect to the average certainty equivalent associated with the Walrasian allocation. This latter statis-

tics is zero for the batch auction and the continuous double auction, because within these protocols all exchanges take place among traders who never accept disadvantageous trades. It is relevant to evaluate the loss of welfare imposed by the presence of a specialist dealer. Because of the bid-ask spread, when a trader with negative excess demand of one unit completes a transaction with a trader with positive excess demand of one unit by going through the dealer, they end up jointly losing some money to the dealer. The ranking for the monetary loss in the average certainty equivalent when the efficient allocation is reached is $0 = \{\text{Loss}_B, \text{Loss}_C\} < \text{Loss}_H < \text{Loss}_D$. This ranking is only partially confirmed from a dynamic point of view in Figure 3, which on the right reports the relative differences between the average certainty equivalent of a protocol and that one of B. (The absolute levels are shown on the left panel.) While B consistently exhibits the highest average certainty equivalent, during the transient phase C switches from having the worst performance to being as good as B. In other words, although C eventually minimizes the loss in the average certainty equivalent, it has the worst performance during the first few trading sessions.

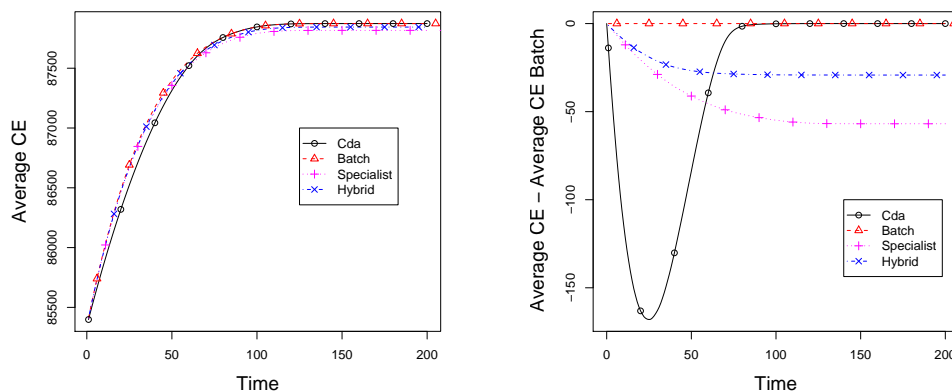


Figure 3: TT: average certainty equivalents (left) and relative differences (right).

Finally, the eight and ninth columns in Table 2 report the standard deviation and the kurtosis of the time series of the closing prices from the first trading day until NT. The ranking with respect to overall volatility of prices is $\sigma_H < \{\sigma_B, \sigma_D\} < \sigma_C$. Most of the time, the protocols generate empirical distributions for prices that are leptokurtic; however, the kurtosis for D and H tends to be much closer to the value associated with a normal distribution. In a different context, LiCalzi and Pellizzari (2003) has already noted the propensity of the continuous double auction to generate non-normal statistics even under zero intelligence trading. From a dynamic point of view, Figure 4 reports the

standard deviations of the closing prices computed over time windows of 20 trading days. The volatility for C is initially very high, but then settles down to levels similar to H.

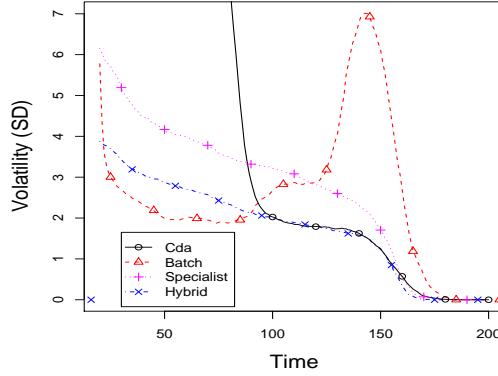


Figure 4: TT: volatility of prices versus time. (Moving window: 20 trading days.)

The volatility for the batch auction increases when approaching the no trade time. This is a well-known consequence of the instability of the k -double auction protocol when the excess demand function has huge flats around the exchange price. We have fixed $k = 1/2$ by assumption, but the instability could be greatly reduced without affecting allocative efficiency by manipulating the value of k . Based on the dynamic comparison, the static ranking of protocols with respect to volatility should be taken with a grain of salt.

4.2 Other behavioral assumptions

The TT- and TT+ assumptions describe behavioral assumptions under which trading is respectively less or more aggressive than TT. Therefore, an analysis of our simulations under $\text{TT}\pm$ provides a test to determine which findings are imputable to our behavioral assumptions. If a ranking under TT is upset when this behavioral assumption is changed, then we cannot claim that this ranking is protocol-driven. Moreover, the analysis provides suggestive evidence for how behavioral assumptions and protocol rules interact.

Table 3 reports summary statistics over the four protocols and the three behavioral assumptions. As before, statistics are computed as averages over 100 different batches of simulations. The columns have the same meaning as in Table 2, but for each protocol there are now three rows — one for each different behavioral assumption. The central row for each protocol repeats for TT the same statistics given in Table 2 for ease of comparison.

We have checked the averages against the medians for the same data and there are no

Prot.	Beh.	Vol	ExcV	NT	Dist	CE	Loss	SD	Kurt
B	TT−	11107.51	1.30%	490.19	0.00	87873.32	0.00%	2.15	36.19
	TT	11251.32	2.61%	148.72	0.00	87873.32	0.00%	4.61	6.90
	TT+	13398.05	22.2%	148.27	0.00	87873.32	0.00%	3.25	16.56
C	TT−	25117.49	129%	401.26	0.00	87873.32	0.00%	18.25	18.06
	TT	28706.28	162%	149.51	0.00	87873.32	0.00%	31.57	3.77
	TT+	36609.01	234%	175.75	0.00	87873.32	0.00%	31.11	4.04
D	TT−	10853.59	0.006%	143.50	0.24	87817.04	0.064%	4.36	0.85
	TT	11090.54	1.57%	142.97	0.11	87816.42	0.065%	4.16	0.74
	TT+	12356.54	12.6%	160.58	0.018	87810.60	0.071%	3.72	1.11
H	TT−	14396.18	31.3%	158.05	0.012	87840.57	0.037%	2.89	-0.11
	TT	16141.92	47.2%	145.06	0.012	87844.09	0.033%	2.85	0.016
	TT+	28017.38	156%	[1500]	0.0093	87843.56	0.034%	0.88	37.5

Table 3: Summary statistics for 100 passes with $n = 1000$.

substantial differences, with two exceptions for the NT criterion under TT−: the average NT is 490.19 for B while the median is 398.50; similarly, the average NT is 401.26 for C while the median is 358.50. Occasionally, B and C under TT− take much longer to attain the no trading situation. We have also made a qualitative comparison between these averages and the data from a sample of individual simulations and we have found no major individual departures. For the special case of H under TT+, in each simulation the market takes much longer than 1500 trading sessions² to attain no trading. This occurs because a few spare traders who initially accumulate a huge trading surplus (by trading with the dealer) end up making a lot of wasteful trades (with their peers) until they dissipate their surplus. Once the surplus is down to zero, convergence takes place as usual. For all other combinations of protocols and behavioral assumptions, convergence is usually attained within 500 sessions. This large discrepancy makes the statistics for H under TT+ inappropriate for comparison, and hence we drop this case from consideration in all of the following discussion.

The major result in this section is that there are no ranking reversals with respect to TT for four of our six performance criteria: volume of trade, distance from the efficient allocation, loss in the average certainty equivalent, and kurtosis. Both the rankings deduced from the table or from the dynamic comparisons (not shown here) are always the same as in TT. Regardless of the exact behavioral assumptions, the qualitative performance of the protocols is analogous.

² We stop the simulations at $t = 1500$ so the statistics for NT is not informative.

Consider for instance volume. The ranking deduced under TT is $\{\text{Vol}_B, \text{Vol}_D\} < \text{Vol}_H < \text{Vol}_C$. Both TT+ and TT- refines this ranking to $\text{Vol}_D < \text{Vol}_B < \text{Vol}_H < \text{Vol}_C$. This suggests that in general the dealership is more likely to avoid wasteful trades than the batch auction. Incidentally, we note that having D in the first place is consistent with the spirit of Tatur (2005, p. 519), where it is shown that “the introduction of a small tax [the dealer’s bid–ask spread] has a first order effect on efficiency”.

Even when protocols rank similarly under different behavioral assumption, Table 3 may provide a second layer of information. Within the same protocol, the impact of more or less aggressive trading surfaces in the comparison of the absolute levels of a performance measure. Consider again volume. Within each protocol, as we move from TT- to TT+ the cumulative traded volume increases. Therefore, more aggressive trading inflates the number of unnecessary trades independently of the trading protocol. On top of this, more aggressive trading in a dealership increases the losses in the average certainty equivalent while reducing the distance of the final allocation from the efficient one.

For the remaining two criteria it is not possible to separate the effect of the protocols from that of the behavioral assumptions, but some clear points emerge anyway. Consider first the time for convergence to no trading. Comparing the statistics about NT for B and C against those for D and H shows that the presence of a dealer significantly reduces this time under the less aggressive trading associated with TT-. Intuitively, the higher efficacy sought by traders under TT- slows down the attainment of efficiency only in the absence of the price-stabilizing influence of the specialist. On the other hand, it is not clear which protocols converge faster in general. It is worth noting that TT tends to improve the time for convergence, presumably because it is more likely to generate the right prices. Once a protocol has reached an allocation sufficiently close to the efficient one, it has to wait until one of the few “inefficient” traders draws a price close to p^* in order to complete a transaction. Meanwhile, under TT- several sessions go away with no trades at all; conversely, under TT+ there are too many unnecessary trades that take place at prices away from p^* . Therefore, on average, convergence under $\text{TT}\pm$ takes longer to be achieved.

Regarding the volatility of prices, the original ranking $\sigma_H < \{\sigma_B, \sigma_D\} < \sigma_C$ under TT has a reversal between H and B under TT- while it is refined to $\sigma_B < \sigma_D$ under TT+. If we exclude H from the comparison, the ranking across the three behavioral assumption is always consistent with $\sigma_B < \sigma_D < \sigma_C$. While the advantages of the batch auction in reducing price dispersion with respect to the continuous double auction were already clear from the TT case alone, this makes a case for the batch auction to be more effective at reducing price volatility than the dealership as well. A second effect, instead, separates

the protocols with a dealer from those without. For B and C, price dispersion is minimum under TT- and maximum under TT; therefore, increasing the aggressiveness of trading has no monotonic effect on price volatility. On the other hand, in the presence of a dealer, the protocols D and H exhibit less volatility as trading gets more aggressive.

4.3 Additional tests of robustness

We have tested the robustness of our model under several different instantiations of its parameters. The only variable with a significant impact on the results is the overall liquidity of the system. When there is no sufficient cash in the system, some transactions that would enhance the allocative efficiency violate budget constraints and cannot be carried out. All the simulations reported in this paper are not affected by issues originating from insufficient liquidity.

In order to provide the reader with a more specific appreciation of this robustness, let us turn to the exemplar parametric configuration given in Table 1. We have run separate checks on each parameter, while keeping the others fixed. For each of these checks, we have obtained summary statistics (computed as averages over batches of 25 different simulations) and compared the rankings associated with them. Within a large range of values for each parameter, the rankings stay almost always unchanged. For instance, initializing $n = 250$ instead of $n = 1000$ produces only one change in the rankings: under TT, the kurtosis for B is now lower than for C. Similarly, no appreciable changes in the rankings emerge if we initialize traders' individual parameters so that exact convergence to the equilibrium price cannot take place because the equilibrium price does not fit on the grid of possible values.

A particularly interesting test of robustness was run on the initial dealer's quotes. As described in Section 3.1, the exemplar configuration assumes an initial bid of 746 and an initial ask of 751. This bid-ask interval is not far from the competitive equilibrium price of 760, raising the legitimate suspicion that this might favorably bias the performance of the specialist-based protocols. Therefore, we ran two batches of 25 simulations each assuming a large variation of up to $\pm 33\%$ in the initial dealers' quotes; more precisely, we assumed an initial bid-ask interval of $[495, 501]$ and $[895, 901]$, respectively. We observed no changes in the final rankings. Clearly, when the initial interval is $[495, 501]$, the initial ask quote vastly underestimates the equilibrium price. Therefore, the dealer is initially obliged to go short on stock and match the strong incoming excess demand at prices unfavorable to her. Nonetheless, her quotes recover sufficiently fast that, when the no trade time is reached, the monetary value of her position is invariably increased; in other words, the dealer's subsequent profits from trading with a bid-ask spread suffice to make up for her initial

losses. A symmetric conclusion holds when the initial quotes overestimate the equilibrium price.

A final test of robustness comes from a comparison of our results with those in Bottazzi et alii (2005). Although it is based on a different set of behavioral assumptions, this paper is closest to ours in the literature. In particular, their simulations depend significantly on a parameter η that represents the share of market (versus limit) orders issued by the agents. Their environment is most similar to ours when $\eta = 0$, so we look at this special case in their paper.

For $\eta = 0$, there are three explicit conclusions that compare directly with our findings. First, Section 5.2.1 reports that the “allocative efficiency is relatively insensitive to the particular ecology of agents” in the market. We have confirmed the insensitivity of the allocative efficiency. Second, Section 5.2.2 reports that the mean absolute deviation for returns under C is greater than the mean absolute deviation under B. We have found the same result for the volatility of prices. Third, Table 6 on p. 30 reports that the kurtosis for returns is higher under B than under C. We have obtained the same result for the kurtosis of the distribution of prices. This latter result, in particular, suggests the conjecture that a substantial difference in the kurtosis may be a “signature” for a difference in the two protocols that is robust to the behavioral assumptions. We leave it to future research to ascertain its validity.

5 Conclusions

The experimental literature has collected ample empirical evidence about trading by human agents in the continuous double auction. As summarized in Smith (1982), the evidence shows that allocations and prices converge rapidly to the competitive equilibrium predictions, even if the informational requirements of this protocol are very simple. Gode and Sunder (1993) argues that the robustness of this conclusion is due to an intrinsic ability of the protocol to guide traders towards the efficient allocation. Our study confirms the allocative efficiency of the continuous double auction under two additional behavioral assumptions and, more importantly, extends Gode and Sunder’s claim to other simple protocols.

Our (computerized) experiments show that there are several simple protocols whose ability to achieve allocative efficiency seems comparable and pretty robust. Therefore, their performance should also be assessed over other relevant dimensions. Under our behavioral assumptions, a direct comparison with the batch auction shows that the continuous double auction is an inferior protocol with respect to volume, time to convergence, and volatility

of prices. (The only exception is time to convergence under the TT- assumption.) This strongly suggests that the experimental literature should give more attention to a comparative study of simultaneous versus sequential protocols; see Section 4.3.1.1 in Madhavan (2000) for a related argument.

We extend the comparison to two alternative simple protocols. The first one is a nondiscretionary form of specialist dealership, in which the rule by which quotes react to transactions is entirely automated and the specialist must adjust prices by one in the direction of the last trade completed. The specialist protocol is equivalent to introducing an additional agent in the market, who in a sense brings more rationality to traders' groping for the efficient allocation. However, note that the specialist dealer is not required to exhibit zero intelligence and must accept all trades that she is proposed. It turns out that following the nondiscretionary rule suffices to produce (modest) gains while keeping her inventory under control. Intuitively, although she may occasionally lose money on some trade, our implementation of the specialist dealership improves her wealth on average. Glosten and Milgrom (1985) prove formally that this holds for a more sophisticated version of specialist dealership in an environment with diversely informed traders.

Under our behavioral assumptions, this specialist protocol and the batch auction vie for the best performance with respect to minimizing wasteful trades, time to convergence, and price dispersion. The main drawback for the specialist protocol is that it drains a tiny amount of wealth from the traders. On the other hand, as this dealer's "profit" might be reallocated to traders at the end of the process, a nondiscretionary dealership remains a natural candidate to investigate in the search for effective protocols to achieve allocative efficiency.

We finally tested a hybrid protocol, that sides the specialist dealer with the continuous double auction giving each trader the option to transact at the most favorable quote available. Under our behavioral assumptions, the hybrid protocol reduces the volatility of prices and the gains of the specialist, but is otherwise inferior to the specialist protocol. The reduction in volatility and in the dealer's gains are easily explained. The specialist starts with initial quotes that may be away from the equilibrium price and must keep a fixed bid-ask spread. Almost all trades initially go through her until her quotes adjust to a level compatible with the equilibrium price. From then on, competition from the traders may occasionally provide better quotes than the dealer's bid-ask spread permits, thereby reducing volatility and her gains.

To conclude, we find that under our behavioral assumptions the four protocols generically converge to the efficient allocation in finite time. An extended comparison over other

performance criteria suggests that the all-round ranking has the batch auction and the dealership vying for the first place while the continuous double auction takes fourth place. The exact ranking for the four protocols depends on the weights given to the performance criteria, although we personally judge the nondiscretionary dealership marginally superior. Finally, we remark that these conclusions hold assuming no informational effects and severely restricting the options for strategic behavior on the part of traders.

References

- [1] N. Audet, T. Gravelle and J. Yang (2002), “Alternative trading systems: Does one shoe fit all?”, working paper 2002-33, Bank of Canada, November.
- [2] G. Bottazzi, G. Dosi, and I. Rebesco (2005), “Institutional architectures and behavioral ecologies in the dynamics of financial markets”, *Journal of Mathematical Economics* **41**, 197-228.
- [3] D. Cliff and J. Bruten (1997), “More than zero intelligence needed for continuous double auction trading”, working paper HPL-97-157, Hewlett Packard Laboratories, Bristol.
- [4] L. Glosten and P. Milgrom (1985), “Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders,” *Journal of Financial Economics* **13**, 71–100.
- [5] D.K. Gode and S. Sunder (1993), “Allocative efficiency of markets with zero intelligence traders: Market as a partial substitute for individual rationality”, *Journal of Political Economy* **101**, 119–137.
- [6] L. Hurwicz, R. Radner and S. Reiter (1975), “A stochastic decentralized resource allocation process: Part I and Part II”, *Econometrica* **43**, 187–222 and 363–394.
- [7] M. LiCalzi and P. Pellizzari (2003), “Fundamentalists clashing over the book: A study of order-driven stock markets”, *Quantitative Finance* **3**, 470–480.
- [8] M. LiCalzi and P. Pellizzari (2006), “The allocative effectiveness of market protocols under intelligent trading”, working paper 134, Department of Applied Mathematics at the University of Venice, May. (Available at <http://ideas.repec.org/p/vnm/wpaper/134.html>.)

- [9] A. Madhavan (2000), “Market microstructure: A survey”, *Journal of Financial Markets* **3**, 205–258.
- [10] M.A. Satterthwaite and S.R. Williams (2002), “The optimality of a simple market mechanism”, *Econometrica* **70**, 1841–1863.
- [11] V. L. Smith (1982), “Microeconomic systems as an experimental science”, *American Economic Review* **72**, 923–955.
- [12] G.J. Stigler (1964), “Public regulation of the securities markets”, *Journal of Business* **37**, 117–142.
- [13] T. Tatur (2005), “On the trade off between deficit and inefficiency and the double auction with a fixed transaction fee”, *Econometrica* **73**, 517–570.
- [14] R. Wilson (1968), “The theory of syndicates”, *Econometrica* **36**, 119–132.